Electronic Journal Publishing: A Reader Version 2.0

Published by INASP, 2001

http://www.inasp.info

7.4 The SciELO project for Latin America and Caribbean: advances and challenges of an emerging model for electronic publishing in developing countries

Abel Laerte Packer
scielo@bireme.br
BIREME/PAHO/WHO
Preliminary version, 30 April 1999

1. Introduction

The SciELO – Scientific Electronic Library Online Project started its operation in March 1997 under a partnership agreement between FAPESP (the Foundation for the Promotion of Science of the State of SaoPaulo), BIREME (the Latin American and Caribbean Center on Health Sciences Information, a specialized center of the Panamerican Health Organization / World Health Organization), and a group of ten Brazilian scientific editors.

A combination of demands, motivations, assumptions and expectations emerged in 1996 making the project a reality:

- The demand of an index to control the scientific literature not covered by international indexes. In 1999, only 12 Latin American and Caribbean (LA & C) journals are included in the Journal Citation Reports published by ISI – Institute for Scientific Information. In the health sciences field, 45 journals are indexed in MEDLINE database, while the LILACS – Latin American and Caribbean Health Sciences Literature database indexes about 500 journal titles. Except for LILACS database in the health area, there are no other Brazilian or LA & C database that carries out a systematic bibliographic control of scientific journals published in the Region. Also there are no established mechanisms or procedures for measuring the usage and impact of the LA & C journals, specially the production of indicators based on citation analysis. This situation is highly negative for the visibility of the scientific production which is not included in international indexes. In particular, governmental institutions that support scientific research and stimulate the communication of its results are interested to overcome this situation.
- The belief that the scientific production from developing countries is being lost due to the lack of accessibility.
- The demand of a methodology for scientific electronic publishing to facilitate the transition from paper to electronic format, involving managerial, technological and economic aspects. Most of the scientific journals of LA & C publishers lack the resources to acquire and/or develop and operate electronic publishing in an efficient way.
- The assumption that the intensive usage of appropriated information technology could contribute to the development of a common methodology with a broad approach on scientific electronic journals, including their preparation, storage, publication, preservation and evaluation of usage and impact.
- The predominance of Internet as an universal information medium raised the expectation that publishing on the net increases the visibility and accessibility of scientific journals.

The project has been evolved around three main specific objectives:

- SciELO Methodology
- development of a methodology for electronic publishing of scientific journals, which could evolve to become a common methodology;
- SciELO national sites application of the SciELO Methodology to operate on the Internet national collections of scientific journals, envisaging to increase their visibility, accessibility and impact;
- SciELO network of sites wide dissemination of the methodology through LA & C countries envisaging the cooperative development of a network of SciELO sites in the Region.

It is expected that the accomplishment of these specific objectives brings a contribution to enhance and renew the scientific communication in Latin America & the Caribbean.

2. The SciELO Project: advances towards a model for electronic publishing

Initially, from March 1997 to May 1998, the SciELO project focused mainly on the development of the SciELO Methodology, which was applied on a pilot basis in the operation of 10 Brazilian journal titles selected from different areas.

The first version of the SciELO Methodology was succesfully developed according to three basic principles:

introduction of electronic publishing in paralel to paper based publishing in order to minimize interferences on publishers traditional production process, as well as editorial idiosyncrasies, practices and policies;
compatibility with international standards and initiatives on electronic scientific communication;
usage of appropriate and affordable information technology envisaging a methodology package, which is easy to transfer, and capable to operate throughout the existing conditions in LA & C countries, including tecnological infrastructure, and human and financial resources.

Beginning in June 1998, the SciELO Project entered a second phase with the main focus centered on the development of national electronic journal libraries, starting with SciELO Brazil, based on the core collection that collaborated in the development of the methodology.

During this second phase, SciELO Brazil is planned to augment sistematically its collection of journals until the middle of the year 2000, when it is expected to stabilize around a collection ranging from 70 to 100 journal titles, covering issues since 1997. In April 1999, SciELO Brazil publishes 29 titles, with more than 2,700 articles in its database, covering different scientific areas, such as Health and Biological Sciences, Social Sciences, Earth and Agricultural Sciences, Physics etc.

SciELO Brazil Internet site http://web.archive.org/web/20040603142438/http://www.scielo.br/ is available free of charge. It is possible to access the entire collection or individual journal titles, as well as to navigate through the issues and articles. It is also possible to search on different bibliographic data, such as author, title words, abstract words, keywords, etc. All articles are available in HTML format, and a small number of journals also publishes in PDF. In addition, the site publishes basic reports on the usage of the site, and a preliminary set of citation analysis based tables. The number of requests for pages has monthly increased. In November 1998, 3,701 requests for pages were registered; in February 1999, there were 5,457 requests for pages; and in April, this number jumped to 9,601 requests. The reports on citations will earn significance when a critical mass of titles and articles are included in the database, which is likely to become a reality by the end of the year 2000, when most of the journals will contribute with four years of data.

During the second semester of 1998, the Chilean National Comission on Science and Technology (CONICyT) started the implementation of the SciELO Chile, in cooperation with SciELO Brazil. The SciELO Methodology was transferred, a group of 7 scientific journals were selected to integrate the SciELO Chile site http://web.archive.org/web/20040603142438/http://www.scielo.cl/ , which is planned to be launched in May 1999.

The successful operation of SciELO Chile will greatly contribute to the dissemination of the SciELO Methodology in other countries. There is already a positive reaction from different countries, specially from the health sciences editors and publishers, whose journals are indexed in the LILACS database. LILACS bibliographic records are produced by a cooperative network of libraries and documentation centers, forming the Latin American and Caribbean System on Health Sciences Information, which is coordinated by BIREME. And BIREME is pushing the SciELO Methodology as the common solution for the transition of health sciences litrature to the electronic format as a part of its regional project called Virtual Health Library.

According to current initiatives, it is expected that SciELO national sites will be operational in 5 countries by the end of 1999, totalling about 40 titles. A Regional SciELO site covering the best public health journals from Latin America is also being implemented.

The establishment and operation of a SciELO site is a complex enterprise. In on side, it requires the

support of national institutions related to scientific communication, such as councils and/or foundations that support scientific research, associations of scientific editors and publishers, scientific societies, libraries, documentation centers, etc. In the other side, it requires the adhesion and commitment of journal editors and publishers. In addition, SciELO operation requires the concourse of skilled managerial and technical human resources, as well as an adequate infrastructure of information technology. SciELO Brazil, for example, is led and financed by FAPESP, a renowned science institution in the country, and by BIREME, a regional center with 30 years of experience on technical cooperation on health sciences information. SciELO Brazil was developed from the very beginning with the active participation of scientific editors and publishers. The ABEC (Brazilian Association of Scientific Editors) early granted SciELO with political support, and contributed to the dissemination of the project. Other national institutions also support the project, such as IBICT (the Brazilian Institute on Scientific and Technical Information). SciELO Chile is being developed along the same lines. It is led by CONICyT, the Chilean national science council, with the cooperation of editors and publishers, and in the health area with the cooperation of the national coordinator center of BIREME.

The institutional configuration and cooperative approach that caracterize the Brazilian and Chilean national projects emerge as a key factor on the future development and dissemination of the SciELO Methodology and the network of SciELO sites. This approach at national level will certainly contribute to strengthen cooperation on scientific communication at the regional level, specially to address common problems related to the promotion and evaluation of the quality, visibility and impact of LA & C scientific journals.

The development of the SciELO project is promoting the rise of a model for scientific electronic publishing in Latin America & the Caribbean – the SciELO Model. Such a model is comprised by two components at the operational level, and a third component at the managerial level. The first is the SciELO Methodology, which may be applied to publish on the Internet a collection of journals or one individual journal. The second component is the SciELO national site, which is intended to include only journals that comply with well-defined standards of scientific communication. In some cases, the second component may be implemented at regional or sub-regional level, when appropriate. The third component refers to the institutional configuration that supports the operation of SciELO sites at national level, which according to the current experience is formed by the alliance of key players in the national scientific communication process.

3. The SciELO Model: challenges and foreseeable solutions

As the SciELO Model emerges, several challenges are posed to its development. They are specially important in three aspects: the technological and methodological treatment of texts, the quality control expressed in terms of definition and application of scietific journal evaluation criteria and policies, and, finally, the economic sustainability.

The treatment of full texts of scientific journals in electronic format does not have an established international standard that is followed by the majority of editors and publishers. The standard format to guarantee full publishing on the Internet is the HTML, a markup language centered on the presentation aspects rather than on the contents and strutuctures of the texts. Therefore, when the final HTML text is sent to the Internet browser, all its bibliographic elements and structure are not formally identified. PDF is another format frequently used to transport and display full texts on the Internet, but its main vocation is to mimic the print version. PDF does not take advantage of the inherent connectivity of the Internet, and it also fails in the possibility of containing extensive marked texts. Making texts redundantly available in HTML and PDF formats is a common practice. However it implies duplication without taking full advantage of searching, tabulating, and connecting text content elements.

In the case of the SciELO Methodology, all texts to be published on the SciELO sites are initially treated by publishers for paper printing. Different journals adopt different technologies and formats to carry out desktop publishing process, including, for example, PageMaker, QuarkPress, Latex and Ventura. Follow on the process, texts in those formats are converted to HTML; the bibliographic elements are marked up; and the texts are loaded into the full text database, from where they are ready to be retrieved by their contents, re-converted to plain HTML, and then transferred to the Internet browser. This process demands intensive manual intervention, and conveys several limitations due to the need of compromising

the text treatment for paper printing and the text preparation for publishing on the Internet. To completely overcome such limitations and implement an integrated text treatment, it is necessary to proceed through a process of full editing and marking up the texts before the process of printing and loading into the full text database. In other words, the markup process preceeds the (hopefully automatic) formatting for paper printing or for electronic display. The foreseeable solution for SciELO Methodology is to use XML (Extensible Markup Language) for preparing electronic texts, from which they can be printed, displayed in a monitor, or processed via other media. However, the introduction of this solution will imply an intervention on the current desktop process the journal publishers are presently using, and it will probably take one to two years to be implemented.

Quality control is becoming an important issue on the development of the SciELO Model. Its methodology may be used to publish any journal regardless of its scientific standard. However, the SciELO national sites are conceived to contain journals that respond to international standards of scientific communication. The problem here is to define adequate criteria and indicators to properly measure scientific communication standards. Note that systematic data on journal circulation and impact are not available, a problem that SciELO Model is expected to solve. There are several formal requirements regarding the presentation of a scientific journal that may be measured, but those elements may be easily accomplished specially under electronic format. The definitive challenge relies on how to measure content quality. SciELO Brazil uses the FAPESP methodology to evaluate and rank scientific journal quality standards. It is a detailed combination of formal and content criteria, whose continued application is higly complex. SciELO Chile utilizes CONICyT criteria to select journals. The development of a set of basic evaluation criteria that stresses recommended practices on scientific communication, and that can easily be implemented is a challenge SciELO is facing in order to establish an efficient, effective, rapid, flexible and transparent way to evaluate journals for entering SciELO national sites. The foreseeable solution is to reach those criteria in the near future, apply them in a flexible way to select journals for SciELO national sites, and then progressively tighten their application in order to induce to a improvement of the performance of the SciELO journals.

Finally, the search for policies for economic sustainability is crucial for the development of the SciELO Model. Its application still needs to be fully funded for more two or three years at least. There are two key issues that directly affect the development of an economic sustainability policy. The first is the fact that most of the paper versions of LA & C journals are not self-financed. This means that the transition to the electronic publishing, without interferences on the paper version production, demands additional financial support. The second is the fact that a key specific objective of the SciELO Model is to increase the visibility and accessibility of LA & C scientific journals on the Internet, reaching new readers nationally and internationally. SciELO Model may apply any of the current practices of delivering paid access of electronic journals on the Internet, but any initiative in the area is likely to be implemented when the SciELO site and its individual journals achieve the necessary visibiltiy in order to expand a potential market interested to purchase online subscriptions. Meanwhile, SciELO Model may contribute to lessen the publication costs of those journals, preparing them for electronic format only.

The strategy is to stress technical cooperation to address and overcome those challenges.